

RESEARCH ARTICLE

The Identification of Circulating MiRNA in Bovine Serum and Their Potential as Novel Biomarkers of Early *Mycobacterium avium* subsp *paratuberculosis* Infection

Damien Farrell¹✉, Ronan G. Shaughnessy¹✉, Louise Britton¹, David E. MacHugh^{2,5}, Bryan Markey¹, Stephen V. Gordon^{1,3,4,5*}

1 UCD School of Veterinary Medicine, University College Dublin, Dublin, Ireland, **2** UCD School of Agriculture and Food Science, University College Dublin, Dublin, Ireland, **3** UCD School of Medicine, University College Dublin, Dublin, Ireland, **4** UCD School of Biomolecular and Biomedical Science, University College Dublin, Dublin, Ireland, **5** UCD Conway Institute, University College Dublin, Dublin, Ireland

✉ These authors contributed equally to this work.

* stephen.gordon@ucd.ie



OPEN ACCESS

Citation: Farrell D, Shaughnessy RG, Britton L, MacHugh DE, Markey B, Gordon SV (2015) The Identification of Circulating MiRNA in Bovine Serum and Their Potential as Novel Biomarkers of Early *Mycobacterium avium* subsp *paratuberculosis* Infection. PLoS ONE 10(7): e0134310. doi:10.1371/journal.pone.0134310

Editor: Pere-Joan Cardona, Fundació Institut d'Investigació en Ciències de la Salut Germans Trias i Pujol, Universitat Autònoma de Barcelona, CIBERES, SPAIN

Received: April 20, 2015

Accepted: July 7, 2015

Published: July 28, 2015

Copyright: © 2015 Farrell et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: Sequencing data are available in the ArrayExpress database (www.ebi.ac.uk/arrayexpress) under accession number E-MTAB-3445.

Funding: This work was supported by the Department of Agriculture Food and the Marine grants RSF11/S/141 (ICONMAP) and EMIDA ERA_NET 11/RD/EMIDA/1 (MycobactDiagnosis). The funders had no role in study design, data

Abstract

Mycobacterium avium subspecies *paratuberculosis* (MAP) is the aetiological agent of Johne's disease (JD), a chronic enteritis in ruminants that causes substantial economic losses to agriculture worldwide. Current diagnostic assays are hampered by low sensitivity and specificity that seriously complicate disease control; a new generation of diagnostic and prognostic assays are therefore urgently needed. Circulating microRNAs (miRNAs) have been shown to have significant potential as novel biomarkers for a range of human diseases, but their potential application in the veterinary sphere has been less well characterised. The aim of this study was therefore to apply RNA-sequencing approaches to serum from an experimental JD infection model as a route to identify novel diagnostic and prognostic miRNA biomarkers. Sera from experimental MAP-challenged calves (n = 6) and age-matched controls (n = 6) were used. We identified a subset of known miRNAs from bovine serum across all samples, with approximately 90 being at potentially functional abundance levels. The majority of known bovine miRNAs displayed multiple isomiRs that differed from the canonical sequences. Thirty novel miRNAs were identified after filtering and were found within sera from all animals tested. No significant differential miRNA expression was detected when comparing sera from MAP-challenged animals to their age-matched controls at six-month's post-infection. However, comparing sera from pre-infection bleeds to six-month's post-infection across all 12 animals did identify increased miR-205 (2-fold) and decreased miR-432 (2-fold) within both challenged and control groups, which suggests changes in circulating miRNA profiles due to ageing or development (P<0.00001). In conclusion our study has identified a range of novel miRNA in bovine serum, and shown the utility of small RNA sequencing approaches to explore the potential of miRNA as novel biomarkers for infectious disease in cattle.

collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

Introduction

Johne's disease is a chronic enteritis of ruminants that is caused by *Mycobacterium avium* subspecies *paratuberculosis* (MAP) infection [1]. Clinical features such as prolonged diarrhoea, progressive wasting and reduced milk yields have obvious implications for animal health and welfare and translate into substantial economic losses, with the dairy industry being particularly vulnerable [2]. In herds where MAP infection is a problem significant resources are therefore expended in reducing the prevalence and transmission of MAP [3].

MAP infection stages can be sequentially classed as silent, subclinical, clinical and advanced [4]. T_H1 immune responses are characteristic of silent infections, with IFN- γ being the hallmark cytokine, and which can effectively contain mycobacterial infection [5]. This cell-mediated response prevails within 85–90% of infected animals and prevents the onset of clinical signs and bacterial transmission. The other 10–15% of animals eventually enter the subclinical phase, where T_H1 responses gradually decline and T_H2 responses become more dominant [4,5]. The T_H2 humoral response offers little protection against disease progression and faecal shedding of MAP commences [6]. At this point, transmission within herds through the faecal-oral route is enabled and becomes more established as shedding increases in line with the onset and progression of the clinical disease. Isolation of infected animals in the early subclinical phase before shedding is established would serve to halt MAP dissemination within herds. However, the ability to accurately identify animals at an early stage of infection that are progressing towards shedding is critical to the success of such a disease control approach and requires a robust prognostic biomarker.

Faecal culture is regarded as the “gold standard” for the diagnosis of MAP infection but requires lengthy incubation periods, and sensitivity can be low in early subclinical infections [7]. Faecal PCR assays are less reliable in the early stages of disease, with one study reporting a detection rate of 4% for low to moderate MAP shedders [8]. IFN- γ based diagnostics have problems of specificity, as the crude MAP antigen preparations used in the whole-blood stimulations show cross-reactivity to antigens from other, environmental mycobacterial species. Similar specificity issues also arise with the available serological assays [9]. In addition, serological assays are of low sensitivity and MAP-specific antibodies are only detected in about 15% of subclinical cases [10]. Overall, it is clear that novel and reliable MAP-specific diagnostic and prognostic assays are required.

Recently, microRNAs (miRNAs) have attracted attention as potential prognostic and diagnostic biomarkers for numerous human pathologies [11]. These short (~22 nt) non-coding RNAs regulate mRNA expression and are predicted to target at least one-third of known mammalian genes [12,13]. Some of the key miRNAs that are known to regulate immune responses to intracellular pathogens have been reported to be upregulated in several cell types in response to mycobacterial infections [14,15]. Thus, specific miRNA expression signatures in tissue samples appear to reflect the underlying host-defence processes in different tissues/cells, and it is possible that different stages of infection have distinct miRNA signatures. Identifying miRNA signatures in the circulation is a more attractive option, as sampling is less invasive and the extracellular miRNAs present are resistant to degradation due to their containment within vesicles or association with proteins [11,16]. Several studies have shown that miRNA profiles in human sera samples of active tuberculosis (TB) patients differ from those of healthy controls [17–21]. It is therefore possible that specific circulating miRNA profiles exist for other mycobacterial infections, including Johne's disease in cattle.

miRNA-sequencing can simultaneously identify and quantify the full repertoire of miRNAs in the circulation at any specific physiological state. A major hindrance with sequencing a cellular body fluids, however, is that their RNA concentrations are below the requirements of

most commercial (small RNA) library preparation kits. Nevertheless, there has been some recent notable successes with small biofluid volumes [22,23], suggesting that this approach may be applicable under similar circumstances.

Our hypothesis in this study was that MAP infection would cause significant changes in the miRNA profile of bovine serum and thus offer a novel diagnostic modality where MAP infection could be disclosed via miRNA profiles. To explore this, we applied RNA sequencing approaches to samples from an experimental MAP infection model. Small RNA-sequencing for miRNA detection on bovine serum has yet to be reported and thus a key aim of this study was to determine whether miRNAs could be sequenced from low serum volumes (≤ 1 ml). In our infection model, 3–6 week old calves were infected with a high MAP-challenge dose, and the course of infection followed over the first 6 months of infection. This early period would typically represent the silent phase of infection, when faecal shedding and any serological response would be negative. Parallel whole blood IFN- γ responses to MAP antigens (PPD-J) provided an immune reference in which to place the miRNA response. A data analysis pipeline was developed for bovine miRNA discovery, and a number of novel miRNAs were identified. Comparison of the miRNA repertoire between infected and control animals allowed an initial assessment of the potential of miRNAs as biomarkers of early silent stage MAP infection.

Materials and Methods

Bacterial cultivation and experimental challenge

Mycobacterium avium subspecies *paratuberculosis* (MAP) strain CIT003 (a gift from Dr Jim O'Mahony, Cork Institute of Technology, Ireland) was cultivated to mid-exponential phase in a modified Middlebrook 7H9 broth consisting of 0.47% Middlebrook 7H9 powder (Becton Dickinson, Dublin, Ireland), 0.05% Casitone (Becton Dickinson, Dublin, Ireland), 0.25% Glycerol (Sigma-Aldrich, Wicklow, Ireland), 10% Middlebrook OADC (Becton Dickinson, Dublin, Ireland) and 2 μ g/ml Mycobactin J (Synbiotics, Lyon, France) at 37°C with rolling at 4 rpm. Mycobacterial CFU numbers were estimated using both the “pelleted wet weight method” [24] and optical densitometry [25]. Based on these quantification approaches, an inoculum dose was prepared to represent approximately 2×10^9 CFU; this was used to infect thirty-five male Holstein-Friesian calves between three to six weeks of age, with the inoculum administered orally on two consecutive days. Retrospective plate counts using Middlebrook 7H11 agar (Becton Dickinson, Dublin, Ireland) supplemented with Mycobactin J revealed the viable challenge dose to be 3.8×10^9 CFU. In parallel to the challenged animals, a control group of twenty calves received a placebo consisting of sterile 7H9 broth. Blood samples were routinely collected from all animals in lithium heparin vacutainers (Becton Dickinson, Dublin, Ireland). At intervals over the first six months of the experiment, cell mediated immunity was investigated using the BOVIGAM IFN- γ release assay (Prionics, Life Technologies) according to the manufacturer's specification, with IFN- γ levels in whole blood samples being quantified following stimulations with MAP-derived PPD-J (22 μ g/ml; a gift from Dr Douwe Bakker, CVI, Netherlands). Sera was prepared by centrifuging clotted blood samples from silicon-coated (red top) vacutainers (Becton Dickinson, Dublin, Ireland) at $1200 \times g$ for 20 mins. The collected sera were immediately stored at -80°C. Sera were routinely screened for screened for MAP-specific antibodies using ELISA (IDEXX). Faecal samples were collected from all animals prior to MAP-challenge and at intervals post-infection. These were cultured for 42 days using the TREK ESP para-JEM system (Thermo Scientific) according to the manufacturer's recommendations. The animal work was approved by the Animal Research Ethics Committee of University College Dublin (AREC-P-12-61-Markey) and licensed by the Irish Government Department of Health and Children (B100-2828).

RNA extraction

1 ml serum samples were incubated with 5 ml of QIAzol Lysis Reagent (Qiagen, Manchester, United Kingdom) for 10 mins at room temperature. RNA was subsequently extracted from the lysate using the miRNeasy Mini Kit (Qiagen, Manchester, United Kingdom) according to the manufacturer's specifications. To precipitate the RNA samples, the 60 μ l eluates were mixed with 1/10 volumes of 3M sodium acetate pH 5.2 (Sigma-Aldrich, Wicklow, Ireland), 1 μ l of glycogen (Invitrogen, Ireland) and 3 volumes of 100% ethanol (Sigma-Aldrich, Wicklow, Ireland) and stored overnight at -80°C . The samples were subsequently centrifuged at $16,000 \times g$ for 25 mins at 4°C . The resulting pellets were washed with 70% ethanol, re-suspended in 6 μ l of nuclease-free water and stored at -80°C .

Small RNA sequencing

Libraries were constructed from 5 μ l of the prepared RNA using the TruSeq Small RNA Sample Preparation Kit (Illumina, Eindhoven, Netherlands) with slight deviations from the manufacturer's protocol. Fifteen cycles of PCR were performed, and the resulting amplicons were purified using an Agencourt AMPure XP kit (Beckman Coulter, Ireland). Briefly, 1.8 \times concentrations of AMPure beads were incubated with the library preparations for 15 mins. The bound PCR products were then washed twice with 80% ethanol, and subsequently re-suspended in 20 μ l of nuclease-free water. The quality and quantity of each library preparation was determined using Agilent DNA 1000 chips (Agilent Technologies, Cork, Ireland) with an Agilent 2100 Bioanalyzer. Each construct was diluted to approximately 20 nM, and randomly assigned to one of two different pools. Electrophoresis using Novex 8% TBE Gels (Invitrogen, Ireland) was used to purify the pools and minimise adapter dimer contamination. The pools were each clustered in two lanes of an Illumina HiSeq 2500 Rapid Run Flow Cell (v1) and sequenced in a SE50bp format using Rapid SBS reagents (sequencing performed at Michigan State University RTSF Genomics Core, Michigan, USA).

Sequencing data is available in the ArrayExpress database (www.ebi.ac.uk/arrayexpress) under accession number E-MTAB-3445.

Identification of known and novel miRNAs

miRDeep2 was used to quantify the reads that aligned to known mature miRNAs and for the prediction of potential novel miRNAs [26]. miRDeep2 uses modules to separate the steps of miRNA quantification and novel prediction. The quantifier module maps to a fasta file of known precursors provided by the user (e.g. miRBase for known mature sequences). The output is a table of read counts and read counts normalised to library size that may be used for expression studies. The core miRDeep.pl module performs prediction and scoring of novel microRNAs using the canonical model of micro RNA biogenesis and by examining the patterns of reads aligned to the reference genome (using the mapper.pl module). Putative miRNAs are scored based on the read stack pattern, stability of the pre-miRNA hairpin and homology to previously identified miRNAs. The expression profile of the novel miRNAs is then also quantified. miRBase version 21 was used for known miRNA mapping.

Before submission to miRDeep2, the adapters were trimmed with cutadapt discarding all reads $< 18\text{nt}$. Known tRNA mappable reads were then identified using Bowtie (only reads with zero mismatches were mapped) and the remaining reads used for analysis. An important aspect of the miRDeep2 results are the performance statistics which give an indication of appropriate score cut-offs by estimating false positives at each score level. Our results showed poor signal-to-noise even at high score levels likely owing to the relatively small percentage of mappable miRNAs in each sample. Therefore we chose a stringent score cut-off of 4

(signal-to-noise = 5.5) for the novel miRNAs and a cut-off of 0 for the known set. An additional reads-based filter was used to remove the remaining low abundance hits remaining from the miRDeep2 analysis. This filter removed all results with reads present in less than 50% of samples and mean normalised read count ≤ 150 . Another algorithm, sRNAbench [27], was also applied to the data mainly to serve as a comparison and for isomiR profiling.

Benchmarking of miRDeep2 and comparison with sRNAbench

To get an indication of read depth required for adequate miRNA discovery we took a single file from our dataset and ran both miRDeep2 and sRNAbench iteratively for increasing random subsets of the reads, ranging from 0.5 to 7 million reads, increasing by 1 million at each run.

Identification of small RNA fractions

To ascertain the degree of contamination in each sample caused by degraded nucleic acid and other small RNA species, we applied a custom Python script to each sample. FastQC was used initially to determine the quality but no reads needed to be removed due to low quality. Corresponding samples from two sequencing lanes were combined together by concatenating the files.

We then applied the following steps for each sample:

1. The 3' adapter was trimmed and quality trimming was performed with $q > 20$ using cutadapt.
2. The cleaned reads were collapsed with the copy number of each unique read retained.
3. Collapsed reads were successively mapped with Bowtie [28] to multiple small RNA bovine annotations according to the category given below. All mapped reads were discarded at each iteration and the unmapped remainder passed on to the next mapping. miRbase (release 21). The Bowtie parameters were set to allow 1 mismatch: **bowtie-f-S-v 1—best-un <remaining><input collapsed fasta>><outfile>**
4. The total percentage of reads mapping to each class was calculated using the original copy numbers.

The following RNA categories were used for mapping: known miRNA found from miR-Deep2, tRNAs from the Genomic tRNA Database [29], rRNAs from RFAM [30], other non-coding RNAs from NONCODE [31], snRNA (downloaded from Ensembl BioMart) and the *Bos taurus* reference genome (UMD 3.1).

Conservation analysis for novel miRNAs

A computational method to find conservation across mammal species was used to aid verification of novel miRNAs. This approach is somewhat similar to the one used by the authors of miRNAmirer [32]. A custom Python script using the pyCogent [33] library was written that utilises the Ensembl Compara programming interface. For each novel miRNA, syntenic alignments from Ensembl (39-way mammal low coverage EPO ensembl 79 [34]) were obtained for a set of 12 species with the most complete assemblies. The following metrics were computed for any aligned syntenic regions in available mammal species:

- sequence identity
- if the seed region was conserved

- if the miRNA is found inside a known gene, the gene type (e.g. protein coding) and whether located inside an exon or intron
- whether there was base pair complementary of the seed in the 3' UTR of any enclosing gene

The results were filtered to produce a list of likely conserved miRNAs (those with $\geq 90\%$ identity in at least one other species) and the remainder considered to be non-conserved.

Differential expression analysis

EdgeR [35] was used to perform pairwise differential expression (DE) analyses for both conditions (experimentally infected and control) at each time point. Only the core miRNAs were included in the DE analysis since the remainder would be missing in too many samples (or too low abundance) to be statistically significant. Only samples with FDR < 0.05 based on Benjamini and Hochberg multiple testing correction [36] and log fold change > 1.5 were considered significant.

Determination of a threshold read count for relevant miRNA selection

The read count threshold used to filter the miRDeep results was arrived at by using the method of Koh et al. [37] to estimate the minimum threshold of biological significance of read counts across replicates. Briefly, this method utilises the observation that the read count distribution of each sample can be interpreted as a mixture of two essentially distinct distributions. The low-abundant count provides noise-rich data and is followed by a fairly even distribution across the log scale of counts, tailing off with sporadic high counts representing the high abundance genes. This latter distribution should be similar between replicates since it represents the more consistently expressed transcripts. By iteratively removing the lowest counts from two replicates and comparing the count distributions using the Komolgorov-Smirnov (KS) test it is possible to estimate the point at which replicates gain a high level of similarity. The read count threshold at which the KS value reaches a minimum is selected since it represents the point at which distributions become similar. This is illustrated for our data in S4 Fig.

The Python scripts used to perform all computational work described here are available at <https://github.com/dmnfarrell/mirnaseq>

Results

Interferon-gamma release assay (IGRA) and selection of animals for sequencing

At the three month and six month intervals, all cattle lacked MAP-specific antibodies in their sera and were negative for faecal shedding of MAP (data not shown). This suggested that the animals were still in the silent phase of infection and consequently, assessing cell mediated immunity was deemed appropriate for measuring infection status. The BOVIGAM IFN- γ release assay was used to measure IFN- γ production in whole blood after stimulation with MAP-purified protein derivative (PPD-J). From the experimental animals, 20 MAP-challenged cattle and 20 control cattle were selected for IFN- γ analysis at two-months and six-months post-infection. At both intervals, control animals showed minimal responsiveness to PPD-J, while IFN- γ responses were variable across the challenged animals (Fig 1). Six challenged cattle that had consistently high IFN- γ production at the 2 and 6 month time points post-infection were selected for miRNA-sequencing (highlighted in red, Fig 1). Six animals were selected at random from the control group to serve as age-matched comparisons. The 0 and 6 month

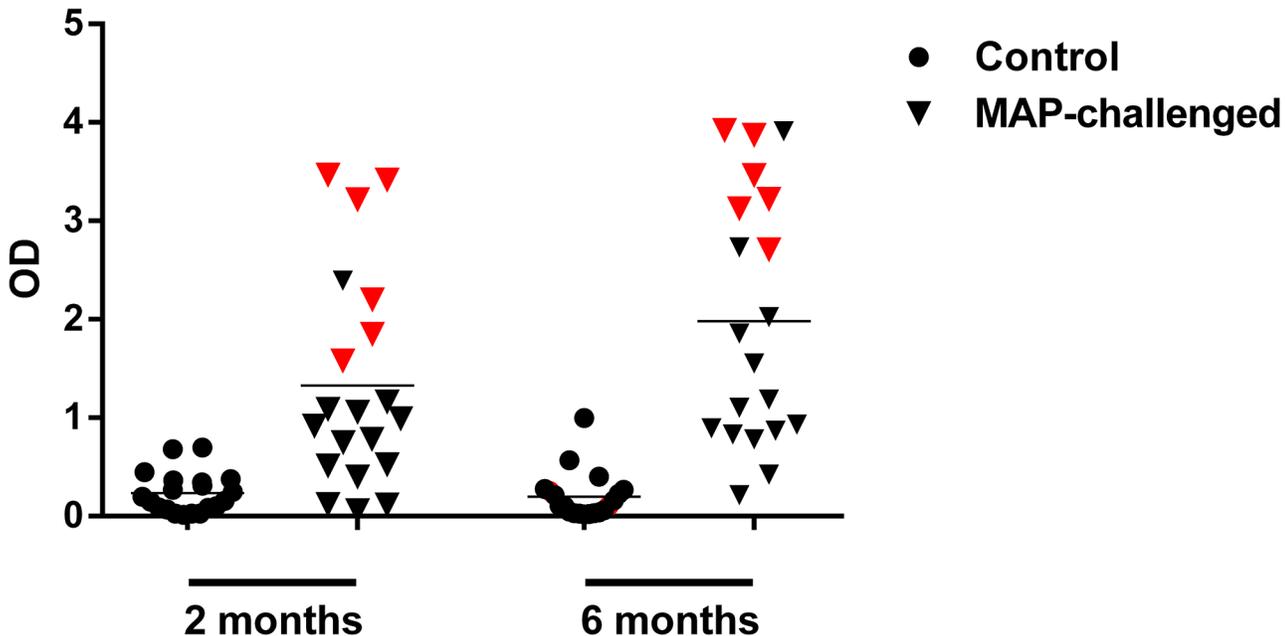


Fig 1. Responses to PPD-J in both sets of animals. IFN- γ production in control (n = 20) and MAP-challenged (n = 20) cattle at 2 months and 6 months post-infection in whole blood in response to PPD-J. Each point represents an individual animal while the red points indicate the six MAP-challenged cattle selected for miRNA-sequencing. The horizontal lines represent mean OD₄₅₀ values for each plot.

doi:10.1371/journal.pone.0134310.g001

time-points were chosen for sequencing analysis to represent two extremes as IFN- γ responses showed maximum differences between groups at these time points.

Identification of small RNAs within bovine serum via sequencing

To the best of our knowledge, this is the first report of successful microRNA sequencing from bovine serum. Thoroughly characterising the serum-specific miRNA repertoire, to identify both known and novel miRNA, will provide valuable insight into circulating miRNAs in cattle. We conducted miRNA-seq on 24 sample libraries resulting in a total of 352,249,513 raw reads. Mapping with Bowtie to *Bos taurus* RNA annotations revealed that only a small percentage of reads in each sample (mean 4.3%) mapped to miRNAs. On average, 58% mapped to tRNA and ~7% to other non-coding RNA databases, while 19% of reads could not be reliably mapped at all (Fig 2A). No alignments were found to the MAP genome. There were significant variations in small RNA composition between individual samples (S1 Fig) but the general trend was consistent. A plot of the typical read length distribution in a sample (S3 Fig) highlights the large amount of degraded small RNA product present in relation to the smaller lengths in the 18–22 range represented by miRNAs.

Read depth analysis and method comparison

To estimate whether the total number of reads were sufficient for miRNA discovery, read depth was compared against the number of miRNAs detected. With the aid of both prediction methods, miRDeep2 and sRNAbench, gradually increasing the read depth correlated with a gradual increase in specific miRNAs. However, the numbers of newly identified miRNAs reached a saturation point at approximately 4 million reads (Fig 2B). Beyond this point, only low abundance genes are newly found by both methods and the majority of these are filtered

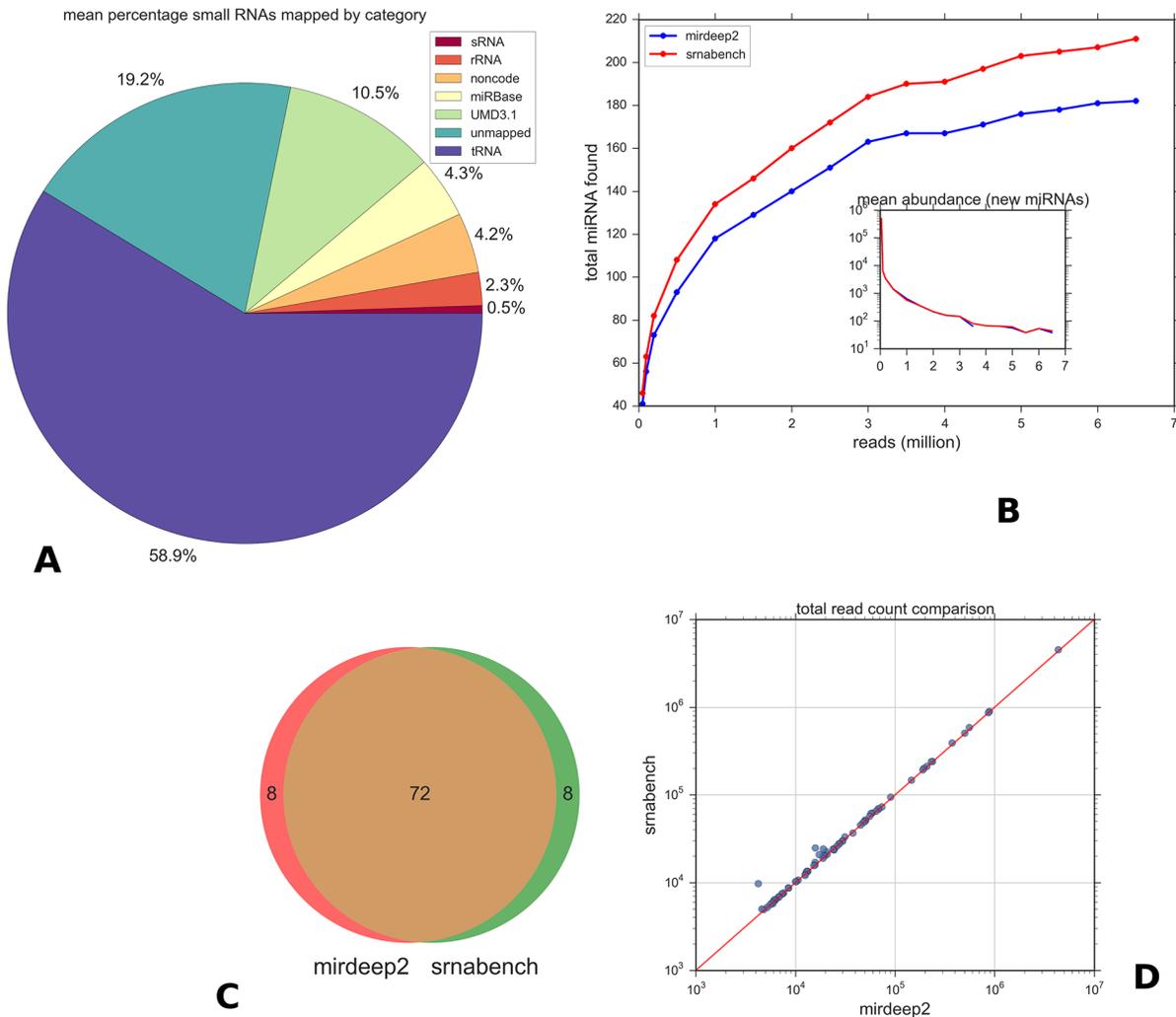


Fig 2. miRNA abundances in miRDeep2 and sRNAbench. A: Proportion of reads mapped to different small RNAs. Values are the average percentage over all samples. Reads matching tRNA dominate. The unmapped percentage is likely due to mismatches in the reads or unannotated ncRNAs. B: Variation of known miRNA discovered with increasing random reads for both methods used in this study. sRNAbench is more sensitive to low abundance genes and produces more hits at a given file size. miRDeep2 was run with no score cut-off. Discovery tails off at 4 million reads in both methods. The inset plot shows the mean abundance of each newly discovered set, illustrating that only low abundance miRNAs are being added after ~3 million mark. C: Overlap between the top 80 miRDeep and sRNAbench results for known miRNAs shows almost identical results. D: Correlation between total read counts determined by both methods for the overlapping miRNAs.

doi:10.1371/journal.pone.0134310.g002

out in actual discovery due to their low levels and low scores. This indicates that the sequencing depth used in this study is more than sufficient for miRNA discovery.

There was a high degree of correlation between sRNAbench and miRDeep2 with regard to the filtered known miRNAs detected (Fig 2C) and their respective total abundance (Fig 2D). In general sRNAbench assigned a slightly higher number of total reads. This is likely due to slight differences in how the canonical miRNA is counted and how length variants and non-templated additions (addition of a nucleotide to the 3' end) are dealt with. Significant differences were found in total read count for only 3 miRNAs (notably bta-miR-99a-5p and bta-miR-99b) as shown by the outlying point in Fig 2D. IsomiR analysis showed that multiple short sub-dominant 3' length variants are seen for these miRNAs by sRNAbench but those reads were not counted by miRDeep2 due to their short length.

For the analyses that follow we used miRDeep2 for expression analyses and novel miRNA discovery, as it provides a good estimate of false positive miRNAs. sRNAbench was used for isomiR profiling because of ease in handling its output.

Known miRNA profiles

Applying the miRDeep2 algorithm to all 24 of our samples and applying the score thresholds and other filters (see [Methods](#) section) produced a set of core miRNAs. The core list consisted of 88 known miRNAs ([Fig 3](#)) of which the top ten most abundant accounted for 85% of the total. In order of decreasing abundance, the top five were miR-486, miR-423-5p, miR-92a, miR-22-3p and miR-191. Our core set thus represents only a fraction of the currently known total miRNA bovine repertoire in miRBase. The majority of our core set have homologs in other mammalian species. Only 13% were from the 3p strand, and there was a roughly linear relationship between abundance (log mean read count across all samples) and frequency (the total number of samples in which the miRNA is found), as shown in [S2 Fig](#). That is, miRNAs at low abundance are observed less frequently, probably due to insufficient depth in a subset of samples. We also compared our core miRNAs to the known human plasma and serum co-fractionate microRNAs contained in the miRandola database [38]. As shown in [Fig 4](#) there is a substantial overlap between the miRNA families with over 70% of our circulating miRNAs being present in the database though they are only a small subset of the total bovine repertoire. There were no significant patterns in global miRNA populations across the two time points.

IsomiR analysis

Sequence variant (isomiR) profiles across all samples were obtained from sRNAbench. The results were filtered to remove all low abundance isomiRs (total reads < 10 across all samples and present in less than 50% of samples) that could be attributable to sequencing or alignment

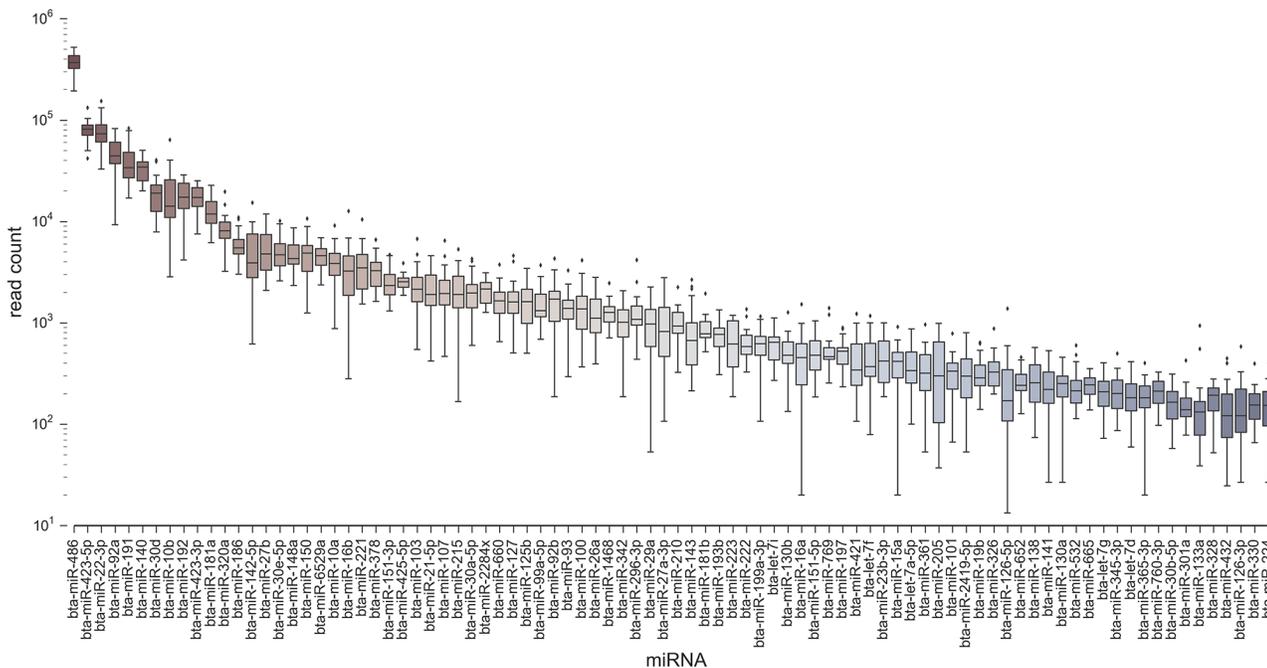


Fig 3. Read count distributions in all samples for top known miRNAs found by miRDeep2.

doi:10.1371/journal.pone.0134310.g003

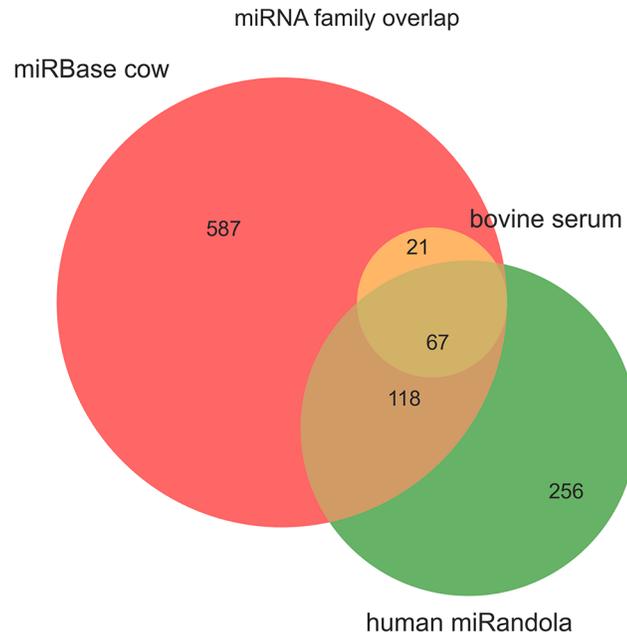


Fig 4. Overlap between core miRNAs in this study and known databases. The known bovine miRnome from miRBase 1.2 is in red. The green set are currently annotated extracellular circulating miRNA human families from the miRandola database. The yellow set represents our core bovine miRNAs.

doi:10.1371/journal.pone.0134310.g004

artefacts. As shown in previous studies [39] for many miRNAs the dominant isomiR was different to the reference miRBase mature sequence. The dominant and reference sequences differed in 51% of cases. Table 1 shows the 20 most significant isomiRs. In some of these cases the reference sequence was not present at all. 99% of miRNAs presented isomiRs with the number of variants seen being highly dependent on the abundance (S5 Fig). Typically isomer length distribution was normal, most obviously when low abundance variants are included.

On average the dominant form accounted for 54% of all reads for a miRNA, but this varied greatly. Variant expression followed two main patterns as seen in other studies [40] those miRNAs with a strong predominant isomiR, such as bta-miR-191 and bta-miR-103; and those miRNAs where there is no predominant isomiR, such as bta-miR-486 and bta-miR-320a. The most common isomeric forms are 3' prime length variants, with the frequencies of all variant classes shown in Fig 5A using the sRNAbench nomenclature. Some 3' trimmed variants as much as 5 or 6 nt shorter than the reference were seen, as shown in Fig 5B, although most of these reflect low abundance forms of abundant miRNAs and may represent mismatches of short degradation products. However in one of these cases, bta-miR-23b-3p, an 18nt 3' length variant constitutes the dominant form with 58% of all reads. This highlights the importance of strict parameters in counting of isomeric variants and how such heterogeneity can affect the consistency and accuracy of abundance in certain cases.

Novel miRNA discovery

After filtering-out low scoring and low frequency results, 30 putative novel miRNAs predicted by miRDeep2 were retained (Table 2). The majority of these are at low abundance levels but present consistently in all samples. Since the most likely candidates are those conserved across related species, we attempted to computationally screen our list by finding evidence of conservation across syntenic regions of other mammal genomes. The approach, (Methods section)

Table 1. Top 20 dominant isomiRs across all samples. Those not marked as 'exact' variants are different from the canonical miRBase sequence. Variants are named using the sRNAbench nomenclature.

name	top isomeric sequence	counts	total	percentage	isomirs	variant
bta-miR-486	TCCTGACTGAGCTGCCCCGA	1339764	4497982	0.30	733	lv3p
bta-miR-22-3p	AAGCTGCCAGTTGAAGAACTGT	665789	866536	0.77	152	lv3p
bta-miR-423-5p	TGAGGGGCAGAGAGCGAGACTTT	409857	890481	0.46	300	exact
bta-miR-92a	TATTGCACTTGTCCCGGCCTGT	325252	582925	0.56	288	exact
bta-miR-191	CAACGGAATCCCAAAGCAGCTG	324529	502395	0.65	186	exact
bta-miR-30d	TGTAAACATCCCCGACTGGAAGCT	131657	238102	0.55	138	exact
bta-miR-25	CATTGCACTTGTCTCGGTCTGA	117960	198545	0.59	104	exact
bta-miR-10b	TACCCTGTAGAACCGAATTTGT	113970	236010	0.48	146	lv3p
bta-miR-140	ACCACAGGGTAGAACCCAGGAC	106449	383774	0.28	354	mv
bta-miR-192	CTGACCTATGAATTGACAGCC	64586	207690	0.31	191	lv3p
bta-miR-423-3p	AGCTCGGTCTGAGGCCCTCAGT	55006	189431	0.29	184	lv5p
bta-miR-181a	AACATTCAACGCTGTCTCGGTGA	45981	144760	0.32	104	lv3p
bta-miR-148a	TCAGTGCACACTACAGAACTTTGT	42423	60311	0.70	56	exact
bta-miR-6529a	GAGAGATCAGAGGCGCAGAGT	39309	50349	0.78	45	exact
bta-miR-30e-5p	TGTAAACATCCTTGACTGGAAGCT	38605	63314	0.61	65	exact
bta-miR-16b	TAGCAGCACGTAAATATTGGCG	38479	47914	0.80	26	mlv3p
bta-miR-186	CAAAGAATTCTCCTTTTGGGCT	31118	71685	0.43	60	exact
bta-miR-27b	TTCACAGTGGCTAAGTTCTG	29769	68708	0.43	46	lv3p
bta-miR-142-5p	CCATAAAGTAGAAAGCACT	27060	67172	0.40	65	mv
bta-miR-375	TTGTTCGTTCCGGCTCGCGTGA	25503	52553	0.49	61	lv5p

doi:10.1371/journal.pone.0134310.t001

uses the Ensembl Compara [34] genome wide comparisons to identify regions of synteny for each of our novel miRNA precursor coordinates and retrieved the aligned sequences for all

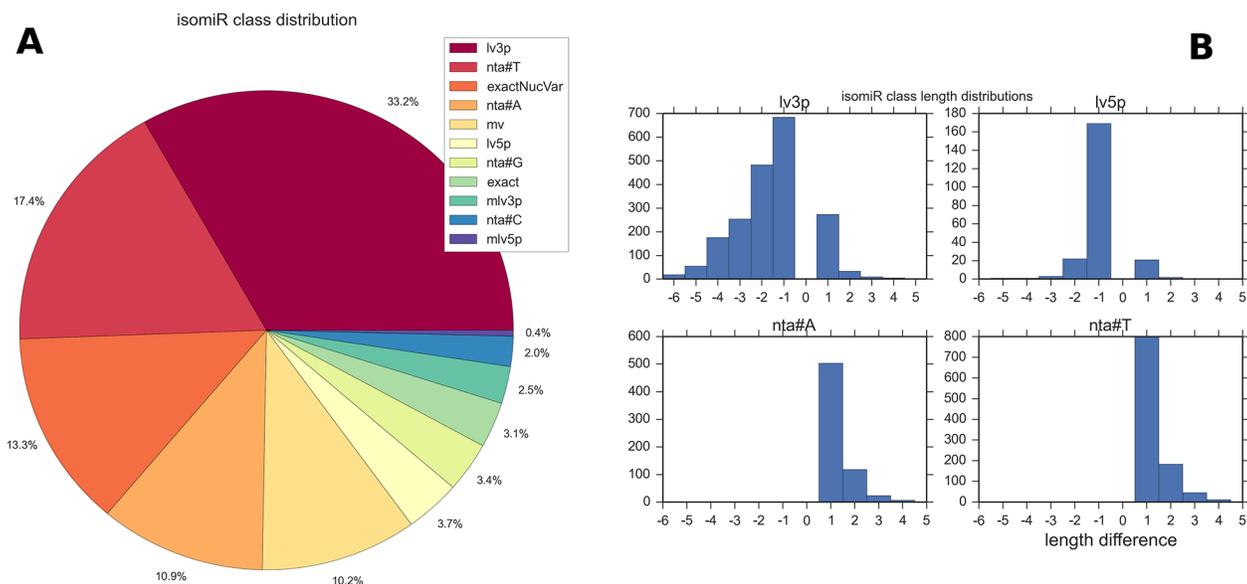


Fig 5. Isomir results. A: The relative frequencies of all isomiR classes including exact matches to the reference. *nta#** denote non-templated additions, *lv3p* and *lv5p* the 5' and 3' length variants and *mv* denotes variants different at both ends. Classification is according to the sRNAbench hierarchical scheme. B: Length distributions for 5' and 3' variants showing the large variation in the latter. (x-axis shows length difference from canonical form). A number of the extreme 3' trimmed sequences are possibly due to erroneous mapping of short reads <18nt.

doi:10.1371/journal.pone.0134310.g005

Table 2. Novel micro RNAs predicted in the data by mirDeep2 (after filtering steps). The 'gene' column denotes a host gene where found with transcription unit (intron or exon). The table is sorted by conservation in mammal species, that is, the number of species with aligned orthologs (aligned column). tr. unit = transcription unit. *max identity is the highest sequence identity to another species.

miRNA	max identity*	aligned	energy	tr. unit	biotype	gene	read count	miRDeep2 score	mirbase seed match	precursor coordinate
20_11724	1	11	-34.6	intron	protein coding	DROSHA	223	133.9	hsa-miR-635	20:42077865..42077925:+
22_13473	1	11	-22.2		miRNA	bta-mir-191	257	133.1	-	22:51543482..51543545:-
15_6321	0.952	10	-49.2				74	79.7	-	15:36372561..36372621:-
3_18437	0.992	10	-105.0		protein coding	SLC35D2	219	136.1	-	3:78601739..78601806:+
3_18843	0.863	9	-42.3				70	71.7	hsa-miR-7977	3:20198831..20198891:-
28_17041	1	9	-24.5				256	5.3	hsa-miR-1469	28:8597865..8597927:-
13_4632	0.99	8	-91.0	intron	protein coding	SNX5	1498	5.7	hsa-miR-4532	13:38543580..38543652:-
10_1684	0.941	7	-22.2				59	4.9	hsa-miR-423-5p	10:13002726..13002777:-
11_2549	0.984	7	-28.1	exon	protein coding	PTGS1	59	33.7	-	11:93244929..93244990:+
13_4213	0.763	5	-36.3	exon	protein coding	SRMS	253	160	-	13:54591461..54591519:+
28_17046	0.873	5	-30.1				3763	5.3	hsa-miR-4766-5p	28:9347067..9347134:-
18_8057	0.673	3	-19.6				177	4.2	hsa-miR-22-3p	18:1864501..1864552:+
9_24948	0.91	2	-18.3				192	5.1	hsa-miR-142-5p	9:86842891..86842958:+
6_21896	0.966	2	-10.3	intron	protein coding	PDCL2	235	120.6	-	6:72695823..72695882:+
5_20820	0.968	2	-28.1	intron	protein coding	ABCC9	71	5.2	hsa-miR-6747-5p	5:88730950..88731013:+
4_20321	1	2	-16.4				186	4.8	hsa-miR-6827-5p	4:115125346..115125387:-
3_18517	0.966	2	-21.6	intron	protein coding	ORC1	1771	1374.1	-	3:94548590..94548649:+
10_1480	0.948	2	-25.1	intron	protein coding	DAAM1	193	38.7	-	10:71809966..71810024:+
3_18032	0.965	2	-31.8				54	5.6	hsa-miR-561-3p	3:2852303..2852360:+
26_16197	0.763	2	-25.9	intron	protein coding	DMBT1	1572	887.3	-	26:42790257..42790316:+
26_16101	0.915	2	-19.6				143	99	hsa-miR-548at-3p	26:26545665..26545724:+
9_25151	0.957	2	-22.9				7363	4.2	hsa-miR-6873-5p	9:27844229..27844299:-
29_18010				intron	protein coding	BRSK2	152	5.9	hsa-let-7i-3p	29:51095161..51095211:-
4_20156				intron	protein coding	FAM3C	197	102.5	-	4:86669291..86669351:-
24_14535			-19.8				227	198.4	-	24:6436128..6436188:+
9_24602				exon	misc_RNA	Metazoa_SRP	491	5.5	hsa-miR-6829-5p	9:2105903..2105987:+

(Continued)

Table 2. (Continued)

miRNA	max identity*	aligned	energy	tr. unit	biotype	gene	read count	miRDeep2 score	mirbase seed match	precursor coordinate
1_987			-23.0				51	4.6	hsa-miR-6890-3p	1:135841882..135841950:-
4_19879							2175	5.1	hsa-miR-515-5p	4:15729..15810:-
25_15966							173	6.2	hsa-miR-132-5p	25:42673013..42673075:-
3_18831							141	95.2	hsa-miR-637	3:19127469..19127551:-

doi:10.1371/journal.pone.0134310.t002

available mammal species. The results could be grouped into three categories: 1) those conserved in at least five species, 2) those conserved only in ruminants (sheep and cattle) and 3) the remainder with no alignments due to lack of synteny with other species. Twelve were found to be contained inside a host protein coding gene, all of which were in introns with the exception of two found in exons. These proportions are consistent with the literature [41] and with our core subset of known miRNAs in which 64% are intergenic, 33% in host gene introns and just two in exons. None of the candidates were found to have targets in the 3' UTR regions of nearby or containing genes.

Of the putative novel miRNAs, several had interesting features and are detailed here. 20_11724 is contained inside the intronic region of the Drosha gene in all species and is >95% conserved in all species. This sequence is annotated as a Gencode predicted human novel miRNA (transcript ID ENSCSAT0000020128) but has not been observed experimentally before. 22_13473 was 99% conserved in all species and was found to be reverse complementary to the bta-miR-191 gene. 26_16197 is present in the *DMBT1* gene along with three other novel miRNAs predicted by miRDeep2 but which did not meet the filtering criteria; these form a possible cluster of up to 4 miRNAs inside this gene. None of these miRNAs have so far been annotated in miRBase for any species. After filtering we only saw one probable false positive in the final list. This was 9_24602, found inside an existing predicted ncRNA Metazoan signal (Ensembl ID ENSBTAG0000048185). Further inspection of the read mappings indicated this may not be a real miRNA because of the inconsistent alignment of reads across the hairpin sequence.

The majority of candidates contained a seed that already exists in a known human miRNA but with the remainder of each sequence not well conserved. Some of the non-conserved hits are likely to be bovine specific. The predicted structures for the six top hits are shown in Fig 6. A csv file containing the aligned sequence, genomic coordinate, sequence identity and folding energy for all miRNAs in each species that could be aligned is given in S3 Table. Read alignments in all samples and scores breakdown from miRDeep2 are given in for all novel miRNAs are given in S1 File.

Differential miRNA expression

Both within group and between group comparisons were used to investigate differential miRNA expression. Stringency was ensured by using a combined FDR of <0.05 and a fold-change of 1.5 as cut-offs for significance. Examining differences between the 0 and 6 month time-point post-MAP infection identified two significantly differentially expressed miRNAs. Specifically at the latter interval, miR-205 was increased (2-fold) while miR-432 was decreased (2-fold). Control animals, however, also displayed similar miR-205 increases (2-fold) and

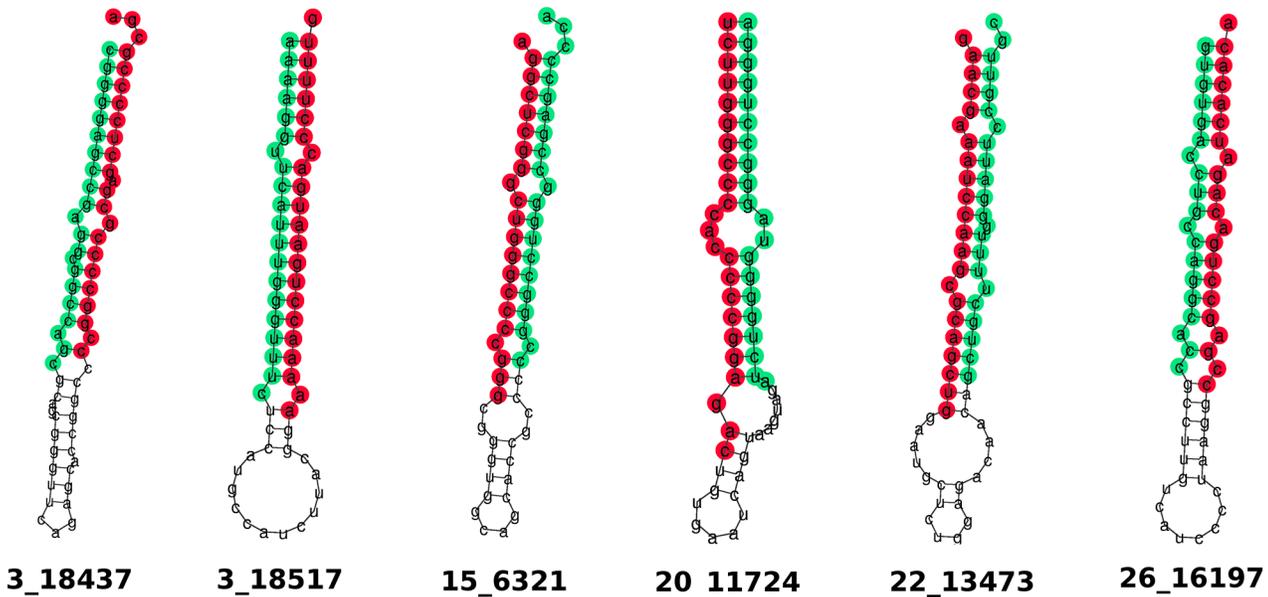


Fig 6. Selected novel miRNAs. Predicted RNA secondary structure of the hairpin for novel micro RNAs. Sequence is coloured according to mature (red) and star (green). Sequences are consensus derived by miRDeep2.

doi:10.1371/journal.pone.0134310.g006

miR-432 decreases (2-fold) at the 6 month time-point but were also accompanied by subtle increases (~1.5 fold) in expression for miR-27a, miR-92b, miR-10b, miR-143 and miR-126-5p (Table 3). Overall, these within group differences appear to be an ageing effect that is independent of infection status as direct comparisons between the infected and control cattle at the six month interval revealed no significantly differentially expressed miRNA.

Subsequently for each group, the reads counts of the eight differentially expressed miRNAs identified during the above pairwise comparisons (miR-10b, miR-126-5p, miR-127, miR-143, miR-205, miR-27a-3p, miR-432 and miR-92b) were plotted to gain further insight into their expression profiles (Fig 7). For all the miRNAs, there were similar expression patterns for the infected and control cattle, but the read counts were somewhat variable at each time-point (Fig 7).

Table 3. Differentially expressed miRNA identified by comparing time-point 0 to time-point 6 months within each group (MAP-infected versus controls).

Gene	Log ₂ Fold Change	P value	FDR
Infected-TP0 vs 6 months			
bta-miR-205	2.14	1.61×10 ⁻¹³	1.39×10 ⁻¹¹
bta-miR-432	-2.23	1.50×10 ⁻¹²	6.45×10 ⁻¹¹
Control-TP0 vs 6 months			
bta-miR-205	2.23	1.54×10 ⁻⁶	9.18×10 ⁻⁵
bta-miR-126-5p	1.67	0.000202	0.002174
bta-miR-143	1.59	1.45×10 ⁻⁵	0.000389
bta-miR-27a-3p	1.55	1.81×10 ⁻⁵	0.000389
bta-miR-92b	1.54	8.55×10 ⁻⁵	0.001050
bta-miR-10b	1.51	3.49×10 ⁻⁵	0.000500
bta-miR-127	-1.78	2.45×10 ⁻⁵	0.000422
bta-miR-432	-2.07	2.14×10 ⁻⁶	9.18×10 ⁻⁵

doi:10.1371/journal.pone.0134310.t003

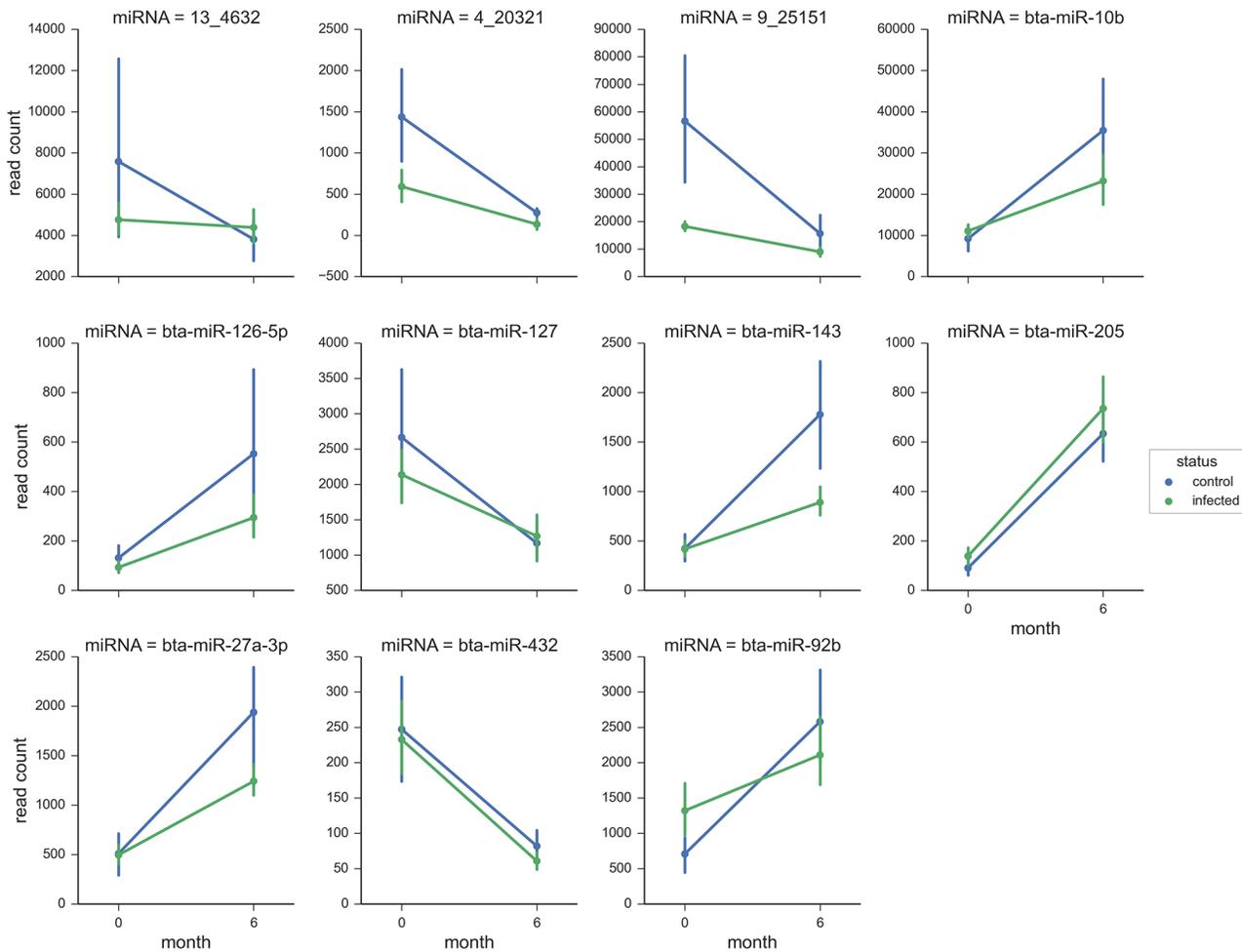


Fig 7. Differential expression results. Normalised read variation for differentially abundant genes from 0–6 months in both control (blue) and infected (green) groups. The bars on each point represent the within group variation for 6 animals (95% confidence interval).

doi:10.1371/journal.pone.0134310.g007

Discussion

Small RNA sequencing is currently the method of choice for identification of the entire repertoire of known and novel miRNAs in a given tissue type or specific disease state [42]. There are however relatively few papers on the application of this technology to circulating miRNA in sera, largely due to very low total RNA concentrations in biofluids. Nevertheless, it has recently been shown that investigation of differential miRNA expression can be achieved with as little as 1 ml of serum [43]. We show here that it is possible to describe circulating miRNA repertoires in bovine serum from low serum volumes, opening up the application of this approach to bovine sera samples and existing biobanks.

It has been reported that large percentages of small RNA reads derived from mammalian sera typically either align to non-miRNA sequences or else cannot be reliably mapped [44–46]. In agreement with these observations, only approximately 5% of our reads mapped to known miRNAs. Notably, a high abundance of reads in each sample corresponded to fragmented tRNAs (25–35 nt), likely arising due to degradation processes that occur within the circulation or as a result of the clotting process during sera preparation. This is, however, not exclusively a serum problem, as Metpally *et al.* also reported a high percentage of fresh human plasma reads

mapping to non-miRNA categories [47]. An obvious concern with low numbers of miRNA-specific reads is whether adequate sequencing depth has been used, and we therefore explored this in our data. By gradually increasing the reads analysed, it was obvious that the numbers of miRNAs identified proportionally increased. However, miRNA discovery appeared to reach a saturation point at approximately 3.5 million reads, a depth above which all our samples had been sequenced at (ranging from 12–29 million reads). While the sequencing depth used may lack the sensitivity to detect rare lowly expressed miRNAs, our depth of coverage was sufficient for the robust identification of potential biomarkers.

We identified approximately 80–100 functionally relevant known miRNAs per sample, with the majority being orthologs of known human circulating miRNAs [48,49]. The top 10 miRNAs accounted for 86% of the total, and miR-486 was the most abundant miRNA across all samples. Interestingly, it has been reported that the IlluminaTruSeq Small RNA Sample Preparation Kit enriches hsa-miR-486 in human plasma libraries yields 50-fold more RNA than other library preparation methods [50], suggesting that some quantitative biases are likely to be apparent within our data. This, however, should not have major implications for differential expression analyses.

Extensive isoform diversity for the majority of miRNAs was shown. Undoubtedly some of this variation is due to technical artefacts such as sequence mismatches of very short reads. However after low abundance filtering, multiple variants are still seen consistently across all samples. These consistent patterns of variant expression provide confidence that such variants are not simply due to degradation products [51]. Distribution of isomiRs from specific tissues is now thought to be non-random and of biological relevance. In serum samples, the picture may be even more complex as multiple subgroups of dominant isomiRs reflecting different tissues may be present. Thus the ubiquity of isomiRs is a potential confounding factor that must be considered in biomarker assays, and even more so for miRNA from biofluids such as blood. The presence of isomiRs could have profound effects on qPCR detection of a miRNA in which several sub-dominant isoforms are actually present that differ sufficiently from the canonical form to cause misdetections.

Various technical factors affect the outcome of miRNA-seq data analyses and it is not always obvious which algorithm is best suited to handling the data; for this reason we chose two approaches, namely sRNAbench and miRDeep2. For instance, sRNAbench was found to be more convenient for isomiR reporting. For several of the lower abundance known miRNAs we found large differences in total read count between sRNAbench and miRDeep2. However, both methods are however highly configurable (for example, the flanking regions within which reads are assigned to a precursor may be varied) and the configuration of multiple profiling parameters could account for significant differences between the two programs.

Tool selection will also have a significant impact on novel miRNA discovery. miRDeep2 was preferred for novel prediction because 1) its scoring model allows transparent screening of false positives and 2) it is designed to be applied to all samples in one run, thus independent samples are easily integrated for the whole study. We found sRNAbench somewhat more cumbersome in this regard. We note that miRDeep2 erroneously missed counting of bta-mir-1246 in the known list and assigned this as a novel miRNA; the reason for this is unclear. Therefore, care must be taken in parsing and filtering the novel candidates provided. However in general miRDeep2 proved to be a powerful ‘off the shelf’ tool for novel discovery once additional filtering was applied. Regardless of the data, it is strongly advised to use at least two tools for initial analysis both for consistency checking and since particular features of one tool may be more appropriate to the data in question [52].

While the cancer and cardiovascular fields already have a significant body of literature dedicated to the potential of miRNAs as biomarkers, to date there have been relatively few

published investigations of circulating miRNA biomarkers applied to infectious bacterial diseases. A few notable exceptions have explored the potential of miRNAs in tuberculosis cases [17–21,53] and as biomarkers of pathology in bovine TB [54]. qPCR-based assays have been the preferred technological approach to identify potential candidates in serum [18,19], but recently miRNA-seq has increased in popularity [20,21] due to the ability to conduct differential expression analysis across the entire “miRNome” of the biofluid.

The lack of robust diagnostic and prognostic assays for MAP infection and resulting disease severely hampers the control and eradication of this infection in cattle. Circulating miRNAs hold promise as a novel diagnostic modality for the disclosure of MAP infection status. The presence of miRNA in peripheral blood is thought to arise due to a number of distinct mechanisms; for example, activated lymphocytes are known to release exosomes containing specific miRNAs [55], while a range of tissue pathologies have distinct circulating miRNA profiles, from fatty liver disease [56] and intracranial aneurysms [57] to prostate cancer [58] and Crohn’s Disease [59]. In this study, we therefore examined MAP-challenged calves displaying early signs of cell-mediated immune responses to MAP, as indicated by IFN- γ responses in infected animals. While calves infected with MAP for only six months would not be expected to show pathology in the gastrointestinal tract, we hypothesised that the cellular responses to infection detected in peripheral blood may give rise to a miRNA profile that would be detectable in serum. A comparison of the serum miRNome of MAP-challenged IFN- γ responders to their age-matched unchallenged controls six months after infection did not identify a significant difference in miRNA expression. This lack of differential miRNA signature in infected animals is likely due to the insidious nature of the infection; at the time-points examined the animals would still be within the silent stage. Future analyses will focus on characterising serum miRNA profiles from cattle that are at a more advanced stage of infection, having developed a serological response to MAP indicative of a T_H1 to T_H2 switch and that have commenced faecal shedding.

Longitudinal differences in miRNA expression were, however, evident and similar within both the challenged and control groups. Comparing the 0 to the six month time-points within each group revealed similar read count profiles for miR-205, miR-10b, miR-92b, miR-432, miR-27a, miR-127, miR-126 and miR-143. All these miRNAs were significantly differentially expressed within the control group but, only miR-205 and miR-432 were identified as significant in the MAP-infected group. These changes in the circulating miRNA profile are likely the result of developmental processes in the calves as all of these miRNAs have been shown to have cellular proliferation and development roles in various mammals [60–67]. Furthermore, in humans, it has been reported that the serum miRNA expression signature (profile) changes with age [68]. Thus, our findings also highlight the critical importance of using age-matched controls in such studies.

In summary we have used deep sequencing of small RNA fractions isolated from bovine serum to explore the utility of circulating miRNAs as potential biomarkers of infection and disease progression in MAP-infection. We have revealed a number of novel circulating miRNA in bovine serum, and identified miRNA whose abundance is linked to bovine development. While no robust miRNA signature of infection could be defined, this may be due to the silent nature of MAP infection at the early time point selected; future studies focussed on animals in the subclinical and clinical phases of infection may prove more revealing.

Supporting Information

S1 Fig. Small RNAs per sample. Proportional mapping by Bowtie to each category of small RNA for each sample in the study. Bars are grouped by pool but not otherwise ordered. (TIF)

S2 Fig. Read count vs. sample frequency. Relationship between log normalised mean normalized read count and sample frequency (number of samples in which each miRNA is found). The majority of high abundance hits are found in over 80% of samples.

(TIF)

S3 Fig. Read length distribution for a representative sample after adapter trimming. The large numbers of reads >30 represent tRNA degradation product and the small peaks around 20 reads represent the miRNA content. Reads <18nt were removed for later miRNA analysis.

(TIF)

S4 Fig. Read count threshold estimation. Thresholds where the read count distribution of a pair biological replicates become similar can be estimated using the Komologorov-Smirnov statistic as a distance measure. This point can be defined as the initial minimum as increasing read count values are removed from each distribution. For our data these values ranged from 50–150 depending on samples. Data shown for here is for three pairs of replicates.

(TIF)

S5 Fig. IsomiR copy number per miRNA versus log total abundance. Unique log IsomiR copy number per miRNA versus log total abundance across all samples shows a linear relationship.

(TIF)

S6 Fig. miRDeep2 dependency of signal to noise and false discovery rate on score.

(TIF)

S7 Fig. Read count distributions in all samples for novel miRNAs found by miRDeep2.

(TIF)

S1 File. PDF files generated by miRDeep2 for all miRNAs discovered in this study.

(ZIP)

S1 Table. Known miRNAs specified by miRDeep2 that meet the filtering criteria described in the paper.

(CSV)

S2 Table. Novel miRNA data (miRDeep2).

(CSV)

S3 Table. Orthologs and sequences of all conserved miRNA in each species found using the Ensembl conservation analysis.

(CSV)

Acknowledgments

This work was supported by the Department of Agriculture Food and the Marine grants RSF11/S/141 (ICONMAP) and EMIDA ERA_NET 11/RD/EMIDA/1 (MycobactDiagnosis). We thank Dr Douwe Bakker (CVI, The Netherlands) for the gift of PPD-J, advice and discussion, Dr Jim O'Mahony (Cork Institute of Technology, Ireland) for the gift of *M. avium* subspecies *paratuberculosis* CIT003, Dr John Browne (UCD) for advice on RNA-seq library preparation and Dr Kevin Carr (Michigan State University, USA) for advice on small RNA sequencing.

Author Contributions

Conceived and designed the experiments: RS LB DEM BM SVG. Performed the experiments: RS LB BM. Analyzed the data: DF RS. Wrote the paper: DF RS SVG.

References

1. Olsen I, Sigurgardottir G, Djonne B (2002) Paratuberculosis with special reference to cattle. A review. *Vet Q* 24: 12–28. PMID: [11918176](#)
2. Cho J, Tauer LW, Schukken YH, Gomez MI, Smith RL, Lu Z, et al. (2012) Economic analysis of *Mycobacterium avium* subspecies paratuberculosis vaccines in dairy herds. *J Dairy Sci* 95: 1855–1872. doi: [10.3168/jds.2011-4787](#) PMID: [22459833](#)
3. Geraghty T, Graham DA, Mullaney P, More SJ (2014) A review of bovine Johne's disease control activities in 6 endemically infected countries. *Prev Vet Med* 116: 1–11. doi: [10.1016/j.prevetmed.2014.06.003](#) PMID: [24997766](#)
4. Gilardoni LR, Paolicchi FA, Mundo SL (2012) Bovine paratuberculosis: a review of the advantages and disadvantages of different diagnostic tests. *Rev Argent Microbiol* 44: 201–215. PMID: [23102470](#)
5. Coussens PM (2001) *Mycobacterium paratuberculosis* and the bovine immune system. *Anim Health Res Rev* 2: 141–161. PMID: [11831436](#)
6. Roussey JA, Steibel J, Coussens PM (2014) Regulatory T cell activity and signs of T cell unresponsiveness in bovine paratuberculosis. *Frontiers in Veterinary Science* 1.
7. Clark DL Jr, Koziczkowski JJ, Radcliff RP, Carlson RA, Ellingson JL (2008) Detection of *Mycobacterium avium* subspecies paratuberculosis: comparing fecal culture versus serum enzyme-linked immunosorbent assay and direct fecal polymerase chain reaction. *J Dairy Sci* 91: 2620–2627. doi: [10.3168/jds.2007-0902](#) PMID: [18565921](#)
8. Wells SJ, Collins MT, Faaberg KS, Wees C, Tavornpanich S, Petrini KR, et al. (2006) Evaluation of a rapid fecal PCR test for detection of *Mycobacterium avium* subsp. paratuberculosis in dairy cattle. *Clin Vaccine Immunol* 13: 1125–1130. PMID: [16928884](#)
9. Scott MC, Bannantine JP, Kaneko Y, Branscum AJ, Whitlock RH, Mori Y, et al. (2010) Absorbed EVE-LISA: a diagnostic test with improved specificity for Johne's disease in cattle. *Foodborne Pathog Dis* 7: 1291–1296. doi: [10.1089/fpd.2010.0541](#) PMID: [20704508](#)
10. Timms VJ, Gehringer MM, Mitchell HM, Daskalopoulos G, Neilan BA (2011) How accurately can we detect *Mycobacterium avium* subsp. paratuberculosis infection? *J Microbiol Methods* 85: 1–8. doi: [10.1016/j.mimet.2011.01.026](#) PMID: [21281678](#)
11. De Guire V, Robitaille R, Tetreault N, Guerin R, Menard C, Bambace N, et al. (2013) Circulating miRNAs as sensitive and specific biomarkers for the diagnosis and monitoring of human diseases: Promises and challenges. *Clin Biochem* 46: 846–860. doi: [10.1016/j.clinbiochem.2013.03.015](#) PMID: [23562576](#)
12. Friedman RC, Farh KK, Burge CB, Bartel DP (2009) Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res* 19: 92–105. doi: [10.1101/gr.082701.108](#) PMID: [18955434](#)
13. Lewis BP, Burge CB, Bartel DP (2005) Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell* 120: 15–20. PMID: [15652477](#)
14. Eulalio A, Schulte L, Vogel J (2012) The mammalian microRNA response to bacterial infections. *RNA Biology* 9: 742–750. doi: [10.4161/ma.20018](#) PMID: [22664920](#)
15. Mehta MD, Liu PT (2014) microRNAs in mycobacterial disease: friend or foe? *Front Genet* 5: 231. doi: [10.3389/fgene.2014.00231](#) PMID: [25076967](#)
16. Turchinovich A, Weiz L, Burwinkel B (2012) Extracellular miRNAs: the mystery of their origin and function. *Trends Biochem Sci* 37: 460–465. doi: [10.1016/j.tibs.2012.08.003](#) PMID: [22944280](#)
17. Abd-El-Fattah AA, Sadik NA, Shaker OG, Aboulftouh ML (2013) Differential microRNAs expression in serum of patients with lung cancer, pulmonary tuberculosis, and pneumonia. *Cell Biochem Biophys* 67: 875–884. doi: [10.1007/s12013-013-9575-y](#) PMID: [23559272](#)
18. Miotto P, Mwangoka G, Valente IC, Norbis L, Sotgiu G, Bosu R, et al. (2013) miRNA signatures in Sera of patients with active pulmonary tuberculosis. *PLoS One* 8: e80149. doi: [10.1371/journal.pone.0080149](#) PMID: [24278252](#)
19. Qi Y, Cui L, Ge Y, Shi Z, Zhao K, Guo X, et al. (2012) Altered serum microRNAs as biomarkers for the early diagnosis of pulmonary tuberculosis infection. *BMC Infect Dis* 12: 384. doi: [10.1186/1471-2334-12-384](#) PMID: [23272999](#)
20. Zhang H, Sun Z, Wei W, Liu Z, Fleming J, Zhang S, et al. (2014) Identification of serum microRNA biomarkers for tuberculosis using RNA-seq. *PLoS One* 9: e88909. doi: [10.1371/journal.pone.0088909](#) PMID: [24586438](#)
21. Zhang X, Guo J, Fan S, Li Y, Wei L, Yang X, et al. (2013) Screening and identification of six serum microRNAs as novel potential combination biomarkers for pulmonary tuberculosis diagnosis. *PLoS One* 8: e81076. doi: [10.1371/journal.pone.0081076](#) PMID: [24349033](#)

22. Williams Z, Ben-Dov IZ, Elias R, Mihailovic A, Brown M, Rosenwaks Z, et al. (2013) Comprehensive profiling of circulating microRNA via small RNA sequencing of cDNA libraries reveals biomarker potential and limitations. *Proc Natl Acad Sci U S A* 110: 4255–4260. doi: [10.1073/pnas.1214046110](https://doi.org/10.1073/pnas.1214046110) PMID: [23440203](https://pubmed.ncbi.nlm.nih.gov/23440203/)
23. Burgos KL, Javaherian A, Bomprezzi R, Ghaffari L, Rhodes S, Courtright A, et al. (2013) Identification of extracellular miRNA in human cerebrospinal fluid by next-generation sequencing. *RNA* 19: 712–722. doi: [10.1261/ma.036863.112](https://doi.org/10.1261/ma.036863.112) PMID: [23525801](https://pubmed.ncbi.nlm.nih.gov/23525801/)
24. Hines ME Jr, Stabel JR, Sweeney RW, Griffin F, Talaat AM, Bakker D, et al. (2007) Experimental challenge models for Johne's disease: a review and proposed international guidelines. *Vet Microbiol* 122: 197–222. PMID: [17467201](https://pubmed.ncbi.nlm.nih.gov/17467201/)
25. Shin SJ, Han JH, Manning EJ, Collins MT (2007) Rapid and reliable method for quantification of *Mycobacterium paratuberculosis* by use of the BACTEC MGIT 960 system. *J Clin Microbiol* 45: 1941–1948. PMID: [17428943](https://pubmed.ncbi.nlm.nih.gov/17428943/)
26. Friedlander MR, Mackowiak SD, Li N, Chen W, Rajewsky N (2012) miRDeep2 accurately identifies known and hundreds of novel microRNA genes in seven animal clades. *Nucleic acids research* 40: 37–52. doi: [10.1093/nar/gkr688](https://doi.org/10.1093/nar/gkr688) PMID: [21911355](https://pubmed.ncbi.nlm.nih.gov/21911355/)
27. Hackenberg M, Rodriguez-Ezpeleta N, Aransay AM (2011) miRanalyzer: an update on the detection and analysis of microRNAs in high-throughput sequencing experiments. *Nucleic acids research* 39: W132–138. doi: [10.1093/nar/gkr247](https://doi.org/10.1093/nar/gkr247) PMID: [21515631](https://pubmed.ncbi.nlm.nih.gov/21515631/)
28. Langmead B (2010) Aligning short sequencing reads with Bowtie. *Curr Protoc Bioinformatics Chapter 11: Unit 11 17*.
29. Chan PP, Lowe TM (2009) GtRNAdb: a database of transfer RNA genes detected in genomic sequence. *Nucleic acids research* 37: D93–97. doi: [10.1093/nar/gkn787](https://doi.org/10.1093/nar/gkn787) PMID: [18984615](https://pubmed.ncbi.nlm.nih.gov/18984615/)
30. Griffiths-Jones S, Bateman A, Marshall M, Khanna A, Eddy SR (2003) Rfam: an RNA family database. *Nucleic acids research* 31: 439–441. PMID: [12520045](https://pubmed.ncbi.nlm.nih.gov/12520045/)
31. Bu D, Yu K, Sun S, Xie C, Skogerbø G, Miao R, et al. (2011) NONCODE v3.0: integrative annotation of long noncoding RNAs. *Nucleic Acids Research*.
32. Artzi S, Kiezun A, Shomron N (2008) miRNAMiner: a tool for homologous microRNA gene search. *BMC bioinformatics* 9: 39. doi: [10.1186/1471-2105-9-39](https://doi.org/10.1186/1471-2105-9-39) PMID: [18215311](https://pubmed.ncbi.nlm.nih.gov/18215311/)
33. Knight R, Maxwell P, Birmingham A, Carnes J, Caporaso JG, Easton BC, et al. (2007) PyCogent: a toolkit for making sense from sequence. *Genome Biol* 8: R171. PMID: [17708774](https://pubmed.ncbi.nlm.nih.gov/17708774/)
34. Vilella AJ, Severin J, Ureta-Vidal A, Heng L, Durbin R, Birney E (2009) EnsemblCompara GeneTrees: Complete, duplication-aware phylogenetic trees in vertebrates. *Genome Res* 19: 327–335. doi: [10.1101/gr.073585.107](https://doi.org/10.1101/gr.073585.107) PMID: [19029536](https://pubmed.ncbi.nlm.nih.gov/19029536/)
35. Robinson MD, McCarthy DJ, Smyth GK (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26: 139–140. doi: [10.1093/bioinformatics/btp616](https://doi.org/10.1093/bioinformatics/btp616) PMID: [19910308](https://pubmed.ncbi.nlm.nih.gov/19910308/)
36. Benjamini Y, Hochberg Y (1995) Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society Series B (Methodological)* 57: 289–300.
37. Koh W, Sheng CT, Tan B, Lee QY, Kuznetsov V, Kiang LS, et al. (2010) Analysis of deep sequencing microRNA expression profile from human embryonic stem cells derived mesenchymal stem cells reveals possible role of let-7 microRNA family in downstream targeting of hepatic nuclear factor 4 alpha. *BMC Genomics* 11 Suppl 1: S6. doi: [10.1186/1471-2164-11-S1-S6](https://doi.org/10.1186/1471-2164-11-S1-S6) PMID: [20158877](https://pubmed.ncbi.nlm.nih.gov/20158877/)
38. Russo F, Di Bella S, Nigita G, Macca V, Laganà A, Giugno R, et al. (2012) miRandola: Extracellular Circulating MicroRNAs Database. *PLoS One* 7: e47786. doi: [10.1371/journal.pone.0047786](https://doi.org/10.1371/journal.pone.0047786) PMID: [23094086](https://pubmed.ncbi.nlm.nih.gov/23094086/)
39. Lee LW, Zhang S, Etheridge A, Ma L, Martin D, Galas D, et al. (2010) Complexity of the microRNA repertoire revealed by next-generation sequencing. *RNA* 16: 2170–2180. doi: [10.1261/ma.2225110](https://doi.org/10.1261/ma.2225110) PMID: [20876832](https://pubmed.ncbi.nlm.nih.gov/20876832/)
40. Timoneda O, Balcells I, Núñez JI, Egea R, Vera G, Castelló A, et al. (2013) miRNA Expression Profile Analysis in Kidney of Different Porcine Breeds. *PLoS One* 8: e55402. doi: [10.1371/journal.pone.0055402](https://doi.org/10.1371/journal.pone.0055402) PMID: [23372853](https://pubmed.ncbi.nlm.nih.gov/23372853/)
41. Kim VN, Han J, Siomi MC (2009) Biogenesis of small RNAs in animals. *Nat Rev Mol Cell Biol* 10: 126–139. doi: [10.1038/nrm2632](https://doi.org/10.1038/nrm2632) PMID: [19165215](https://pubmed.ncbi.nlm.nih.gov/19165215/)
42. Pritchard CC, Cheng HH, Tewari M (2012) MicroRNA profiling: approaches and considerations. *Nat Rev Genet* 13: 358–369. doi: [10.1038/nrg3198](https://doi.org/10.1038/nrg3198) PMID: [22510765](https://pubmed.ncbi.nlm.nih.gov/22510765/)
43. Burgos K, Malenica I, Metpally R, Courtright A, Rakela B, Beach T, et al. (2014) Profiles of extracellular miRNA in cerebrospinal fluid and serum from patients with Alzheimer's and Parkinson's diseases

- correlate with disease status and features of pathology. *PLoS One* 9: e94839. doi: [10.1371/journal.pone.0094839](https://doi.org/10.1371/journal.pone.0094839) PMID: [24797360](https://pubmed.ncbi.nlm.nih.gov/24797360/)
44. Dhahbi JM, Spindler SR, Atamna H, Yamakawa A, Boffelli D, Mote P, et al. (2013) 5' tRNA halves are present as abundant complexes in serum, concentrated in blood cells, and modulated by aging and calorie restriction. *BMC Genomics* 14: 298. doi: [10.1186/1471-2164-14-298](https://doi.org/10.1186/1471-2164-14-298) PMID: [23638709](https://pubmed.ncbi.nlm.nih.gov/23638709/)
 45. Juan L, Tong HL, Zhang P, Guo G, Wang Z, Wen X, et al. (2014) Identification and characterization of novel serum microRNA candidates from deep sequencing in cervical cancer patients. *Sci Rep* 4: 6277. doi: [10.1038/srep06277](https://doi.org/10.1038/srep06277) PMID: [25182173](https://pubmed.ncbi.nlm.nih.gov/25182173/)
 46. Krauskopf J, Caiment F, Claessen SM, Johnson KJ, Warner RL, Schomaker SJ, et al. (2015) Application of high-throughput sequencing to circulating microRNAs reveals novel biomarkers for drug-induced liver injury. *Toxicol Sci* 143: 268–276. doi: [10.1093/toxsci/kfu232](https://doi.org/10.1093/toxsci/kfu232) PMID: [25359176](https://pubmed.ncbi.nlm.nih.gov/25359176/)
 47. Metpally RP, Nasser S, Malenica I, Courtright A, Carlson E, Ghaffari L, et al. (2013) Comparison of Analysis Tools for miRNA High Throughput Sequencing Using Nerve Crush as a Model. *Front Genet* 4: 20. doi: [10.3389/fgene.2013.00020](https://doi.org/10.3389/fgene.2013.00020) PMID: [23459507](https://pubmed.ncbi.nlm.nih.gov/23459507/)
 48. Blondal T, Jensby Nielsen S, Baker A, Andreassen D, Mouritzen P, Wrang Teillum M, et al. (2013) Assessing sample and miRNA profile quality in serum and plasma or other biofluids. *Methods* 59: S1–6. doi: [10.1016/j.ymeth.2012.09.015](https://doi.org/10.1016/j.ymeth.2012.09.015) PMID: [23036329](https://pubmed.ncbi.nlm.nih.gov/23036329/)
 49. Pritchard CC, Kroh E, Wood B, Arroyo JD, Dougherty KJ, Miyajima MM, et al. (2012) Blood cell origin of circulating microRNAs: a cautionary note for cancer biomarker studies. *Cancer Prev Res (Phila)* 5: 492–497.
 50. Huang X, Yuan T, Tschannen M, Sun Z, Jacob H, Du M, et al. (2013) Characterization of human plasma-derived exosomal RNAs by deep sequencing. *BMC Genomics* 14: 319. doi: [10.1186/1471-2164-14-319](https://doi.org/10.1186/1471-2164-14-319) PMID: [23663360](https://pubmed.ncbi.nlm.nih.gov/23663360/)
 51. Chugh P, Dittmer DP (2012) Potential Pitfalls in microRNA Profiling. *Wiley interdisciplinary reviews RNA* 3: 601–616. doi: [10.1002/wrna.1120](https://doi.org/10.1002/wrna.1120) PMID: [22566380](https://pubmed.ncbi.nlm.nih.gov/22566380/)
 52. Li Y, Zhang Z, Liu F, Vongsangnak W, Jing Q, Shen B (2012) Performance comparison and evaluation of software tools for microRNA deep-sequencing data analysis. *Nucleic Acids Research* 40: 4298–4305. doi: [10.1093/nar/gks043](https://doi.org/10.1093/nar/gks043) PMID: [22287634](https://pubmed.ncbi.nlm.nih.gov/22287634/)
 53. Fu Y, Yi Z, Wu X, Li J, Xu F (2011) Circulating microRNAs in patients with active pulmonary tuberculosis. *J Clin Microbiol* 49: 4246–4251. doi: [10.1128/JCM.05459-11](https://doi.org/10.1128/JCM.05459-11) PMID: [21998423](https://pubmed.ncbi.nlm.nih.gov/21998423/)
 54. Golby P, Villarreal-Ramos B, Dean G, Jones GJ, Vordermeier M (2014) MicroRNA expression profiling of PPD-B stimulated PBMC from *M. bovis*-challenged unvaccinated and BCG vaccinated cattle. *Vaccine* 32: 5839–5844. doi: [10.1016/j.vaccine.2014.07.034](https://doi.org/10.1016/j.vaccine.2014.07.034) PMID: [25107461](https://pubmed.ncbi.nlm.nih.gov/25107461/)
 55. de Candia P, Torri A, Pagani M, Abrignani S (2014) Serum microRNAs as Biomarkers of Human Lymphocyte Activation in Health and Disease. *Front Immunol* 5: 43. doi: [10.3389/fimmu.2014.00043](https://doi.org/10.3389/fimmu.2014.00043) PMID: [24575093](https://pubmed.ncbi.nlm.nih.gov/24575093/)
 56. Pirola CJ, Fernandez Gianotti T, Castano GO, Mallardi P, San Martino J, Mora Gonzalez Lopez Ledesma M, et al. (2015) Circulating microRNA signature in non-alcoholic fatty liver disease: from serum non-coding RNAs to liver histology and disease pathogenesis. *Gut* 64: 800–812. doi: [10.1136/gutjnl-2014-306996](https://doi.org/10.1136/gutjnl-2014-306996) PMID: [24973316](https://pubmed.ncbi.nlm.nih.gov/24973316/)
 57. Li P, Zhang Q, Wu X, Yang X, Zhang Y, Li Y, et al. (2014) Circulating microRNAs serve as novel biological markers for intracranial aneurysms. *J Am Heart Assoc* 3: e000972. doi: [10.1161/JAHA.114.000972](https://doi.org/10.1161/JAHA.114.000972) PMID: [25249297](https://pubmed.ncbi.nlm.nih.gov/25249297/)
 58. Sita-Lumsden A, Dart DA, Waxman J, Bevan CL (2013) Circulating microRNAs as potential new biomarkers for prostate cancer. *Br J Cancer* 108: 1925–1930. doi: [10.1038/bjc.2013.192](https://doi.org/10.1038/bjc.2013.192) PMID: [23632485](https://pubmed.ncbi.nlm.nih.gov/23632485/)
 59. Schaefer JS, Attumi T, Opekun AR, Abraham B, Hou J, Shelby H, et al. (2015) MicroRNA signatures differentiate Crohn's disease from ulcerative colitis. *BMC Immunology* 16: 5. doi: [10.1186/s12865-015-0069-0](https://doi.org/10.1186/s12865-015-0069-0) PMID: [25886994](https://pubmed.ncbi.nlm.nih.gov/25886994/)
 60. Bhaskaran M, Wang Y, Zhang H, Weng T, Baviskar P, Guo Y, et al. (2009) MicroRNA-127 modulates fetal lung development. *Physiol Genomics* 37: 268–278. doi: [10.1152/physiolgenomics.90268.2008](https://doi.org/10.1152/physiolgenomics.90268.2008) PMID: [19439715](https://pubmed.ncbi.nlm.nih.gov/19439715/)
 61. Lund AH (2010) miR-10 in development and cancer. *Cell Death Differ* 17: 209–214. doi: [10.1038/cdd.2009.58](https://doi.org/10.1038/cdd.2009.58) PMID: [19461655](https://pubmed.ncbi.nlm.nih.gov/19461655/)
 62. Qian NS, Liu WH, Lv WP, Xiang X, Su M, Raut V, et al. (2013) Upregulated microRNA-92b regulates the differentiation and proliferation of EpCAM-positive fetal liver cells by targeting C/EBP β s. *PLoS One* 8: e68004. doi: [10.1371/journal.pone.0068004](https://doi.org/10.1371/journal.pone.0068004) PMID: [23936298](https://pubmed.ncbi.nlm.nih.gov/23936298/)

63. Guo D, Li Q, Lv Q, Wei Q, Cao S, Gu J (2014) MiR-27a targets sFRP1 in hFOB cells to regulate proliferation, apoptosis and differentiation. PLoS One 9: e91354. doi: [10.1371/journal.pone.0091354](https://doi.org/10.1371/journal.pone.0091354) PMID: [24625818](https://pubmed.ncbi.nlm.nih.gov/24625818/)
64. Rihani A, Van Goethem A, Ongenaert M, De Brouwer S, Volders PJ, Agarwal S, et al. (2015) Genome wide expression profiling of p53 regulated miRNAs in neuroblastoma. Sci Rep 5: 9027. doi: [10.1038/srep09027](https://doi.org/10.1038/srep09027) PMID: [25762502](https://pubmed.ncbi.nlm.nih.gov/25762502/)
65. Vosgha H, Salajegheh A, Smith RA, Lam AK (2014) The important roles of miR-205 in normal physiology, cancers and as a potential therapeutic target. Curr Cancer Drug Targets 14: 621–637. PMID: [25308719](https://pubmed.ncbi.nlm.nih.gov/25308719/)
66. Wang W, Wu S, Shi Y, Miao Y, Luo X, Ji M, et al. (2015) c-MYB regulates cell growth and DNA damage repair through modulating MiR-143. FEBS Lett 589: 555–564. doi: [10.1016/j.febslet.2015.01.012](https://doi.org/10.1016/j.febslet.2015.01.012) PMID: [25616133](https://pubmed.ncbi.nlm.nih.gov/25616133/)
67. Zhao C, Li Y, Zhang M, Yang Y, Chang L (2015) miR-126 inhibits cell proliferation and induces cell apoptosis of hepatocellular carcinoma cells partially by targeting Sox2. Hum Cell.
68. Zhang H, Yang H, Zhang C, Jing Y, Wang C, Liu C, et al. (2015) Investigation of microRNA expression in human serum during the aging process. J Gerontol A Biol Sci Med Sci 70: 102–109. doi: [10.1093/gerona/glu145](https://doi.org/10.1093/gerona/glu145) PMID: [25165030](https://pubmed.ncbi.nlm.nih.gov/25165030/)